Regulatory Toxicology and Pharmacology 62 (2012) 278–291

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Hypothesis-Based Weight-of-Evidence evaluation of methanol as a human carcinogen

Lisa A. Bailey [a], Robyn L. Prueitt [b], Lorenz R. Rhomberg [a,*]

[a] Gradient, 20 University Road, Cambridge, MA 02138, USA
[b] Gradient, Suite 803, 600 Stewart Street, Seattle, WA 98101, USA

ARTICLE INFO

Article history:
Received 27 July 2011
Available online 7 December 2011

Keywords:
Methanol
Formaldehyde
Lymphoma
Hazard identification
Human relevance
Risk assessment
Mode of action

ABSTRACT

Recent scientific debate has focused on the potential for exposure to methanol to cause lymphomas in humans. The concern stems from a few animal studies reporting an association, although evidence suggests the studies may have been confounded by chronic respiratory infection. Although the toxicological evidence for methanol carcinogenesis is weak, two modes of action have been put forth, one involving metabolism of methanol to formaldehyde, followed by formaldehyde induction of lymphoma, and another involving oxidative stress caused by hydrogen peroxide release during catalase-induced metabolism of methanol to formaldehyde. In this article, we apply our Hypothesis-Based Weight-of-Evidence (HBWoE) approach to evaluate the evidence regarding methanol exposure and lymphoma, attending to how human, animal, and mode-of-action results inform one another, tracing the logic of inference within and across all studies, and articulating how one could account for the suite of available observations. Upon comparison of alternative proposals regarding what causal processes may have led to the array of observations as we see them, we conclude that the apparent association between methanol exposure and lymphoma in some animal studies is weak and strains biological plausibility, and is better interpreted as due to confounding or to a mechanism not relevant in humans.

1. Introduction and background

Methanol is naturally produced in small amounts in the human body as a result of metabolism and is also found in many fruits, vegetables, and alcoholic beverages (US EPA, 2009). Methanol also has many commercial uses, as it is a building block for many chemical products, including those used in construction, housing and automotive industries, and consumer products such as varnishes, shellacs, paints, windshield washer fluids, and antifreeze (US EPA, 2009). The demand for methanol has increased in recent years, and is expected to continue to increase, because of its use in biodiesel fuels (US EPA, 2009).

There are no epidemiology studies of methanol carcinogenicity in the published literature. The only human studies that exist are neurological case reports and occupational studies. Recently, however, several animal studies (Apaja, 1980, NEDO, 1985a,b and Soffritti et al., 2002, all cited in US EPA, 2009) have been put forth as supporting potential carcinogenic effects of methanol (US EPA, 2009). Two of these studies (Apaja, 1980, as cited in US EPA, 2009; Soffritti et al., 2002), which reported increased incidence of lymphoma in rodents, may have been confounded by chronic respiratory infection in the animals. The other two studies (NEDO, 1985a,b, as cited in US EPA, 2009), which are methodologically more robust, reported no increased tumor incidence in the animals. The only reported carcinogenic effects from these studies are from US EPA's reanalysis (US EPA, 2009) of the NEDO (1985b) rat study; US EPA reported an increased incidence of pheochromocytomas in high-dose females from reanalysis of the NEDO (1985b) data (US EPA, 2009).

Despite the questions about whether methanol carcinogenesis has been reliably observed in animals, US EPA has proposed two modes of action for methanol carcinogenesis in the bioassays: (1) metabolism of methanol to formaldehyde, followed by formaldehyde induction of lymphoma via one of the modes of action recently proposed in US EPA's draft toxicological review of formaldehyde (US EPA, 2010); and (2) induction of carcinogenesis via oxidative stress caused by hydrogen peroxide release during catalase-induced metabolism of methanol to formaldehyde. The application of the bioassay results and mode-of-action hypotheses as evidence for potential human carcinogenicity of methanol rests on: (1) accepting that the apparently positive bioassays do indeed show increases in tumors that are attributable to methanol, (2) accounting at least hypothetically for why other animal studies do not share these effects, and (3) arguing that humans and the responding animals do share the underlying mode of carcinogenic action that is proposed as the basis for the animal response.

* Corresponding author. Fax: +1 617 395 5001.
E-mail addresses: lbailey@gradientcorp.com (L.A. Bailey), rprueitt@gradientcorp.com (R.L. Prueitt), lrhomberg@gradientcorp.com (L.R. Rhomberg).

0273-2300/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2011.11.015

The current understanding of methanol metabolism (slow half-life for metabolism to formaldehyde – 3 h) and formaldehyde metabolism (fast half-life for metabolism to formic acid – 1 min) raises the issue of whether it is plausible to propose that metabolic generation of formaldehyde is responsible for lymphomas in the animal studies. Moreover, this proposed mechanism inherits all the questions about the biological plausibility of formaldehyde as a cause of lymphohematopoietic cancers, as we have reviewed elsewhere (Rhomberg et al., 2011). With regard to the proposed mode of action involving oxidative stress induction by catalase activity from methanol metabolism, it is notable that catalase is not involved in human metabolism of methanol (only in animal metabolism of methanol), raising the issue of whether the observed phenomena in animals, if they are interpreted as being caused by oxidative stress, are relevant to humans.

The difficulty of naming a satisfying mode of action for generating tumors in the animal studies, coupled with the appearance of elevated tumors only in studies with marked methodological concerns, raises the issue of whether one should consider the animal bioassay results to be observation of an effect actually caused by methanol exposure. The alternative is to ascribe the appearance of apparently elevated tumor rates in some studies to methodological faults in those studies, although this alternative explanation is also hypothetical and proposed after the fact to account for the otherwise incongruous appearance of the tumors. If, despite these difficulties, one proposes that the positive animal studies do indeed constitute observation of methanol's ability to cause tumors in these systems, then to apply this to human risk evaluation raises the question of what the responding animals have in common with the target human population that would make the latter also subject to elevated risk.

In short, one is faced with a contradiction between the apparent (though not certainly causal) association of lymphoma with methanol exposure in at least some animal studies and the apparent (though unproven) implausibility of such a causal effect in view of contradictory results in other bioassays and of current biological understanding of possible mechanisms that would be relevant to human exposures. The apparent contradiction can be reconciled in one of two ways: (1) by accepting that carcinogenic risks in animals are actually increased and positing that the biological implausibility of such increases is somehow mistaken – that is, since the effect appears, it must have a possible causal explanation, and this explanation may find a parallel in humans; or (2) by concluding that doubts about possible mechanisms have merit, and the apparent association of methanol and lymphoma reported in a few animal studies does not in fact indicate a causal connection (and that those studies showing lack of effect are indeed the ones to be taken at face value) – that is, the appearance of some apparent associations is in fact accounted for by chance or by shortcomings, which, according to this view, should be deemed false positive results.

We would argue that the process of evaluating the weight of evidence for human carcinogenesis from methanol largely comes down to evaluating the comparative degree to which each of these alternative accounts is supported by reference to scientific evidence. The weighing should attend to observations that are consistent with and expected from each proffered explanation and its hypothesized underlying causes, but it should also note observations that are discordant. In doing so, one should examine potential reasoning for why these discordant observations should not be regarded as definitively refuting – i.e., how they could, in fact, be consistent with the hypothesis being evaluated. The two competing accounts present different and essentially opposite proposals regarding which observations are to be regarded as revealing true, generalizable phenomena and which ones are leading us astray if we take them at face value.

In the present paper, we evaluate the scientific data relevant to the potential causal association between exposure to methanol and increased risk of lymphoma in humans using the criteria and structured approach to weight-of-evidence examination (Hypothesis-Based Weight-of-Evidence or HBWoE) that we have developed and applied elsewhere (Rhomberg et al., 2010, 2011; Prueitt et al., 2011). We believe that our approach to evaluating and weighing all of the evidence will help to address the concerns and "Roadmap" recommendations expressed by the National Academy of Sciences (NAS) report on US EPA's formaldehyde assessment (NRC, 2011). For example, NAS has recommended a risk assessment approach that moves away from a checklist of steps, and involves a focus on inferential and integrative analysis of hazard identification and dose–response. Our approach evaluates the evidence regarding exposure and disease, attending to how human, animal, and mode-of-action results inform one another by explicitly attempting to trace the logic of inference within and across all studies. It entails articulating how hypothesized causes could account for the suite of available observations – not only noting findings that are consistent with the hypotheses, but also proposing how discordant findings are to be explained. Finally, our approach compares alternative accounts regarding what causal processes may have led to the whole array of available observations, with weight-of-evidence preference being given to those accounts that (when compared to competing accounts) provide more explanatory power by tying different observations together under common causes with less recourse to assumptions, especially assumptions unsupported by evidence that are aimed at reconciling otherwise discordant results.

The NAS report (NRC, 2011) discusses how an evidence-based approach to causal inference should incorporate the Hill criteria (Hill, 1965) for evidence evaluation, but also notes that "the criteria are not rigid and are not applied in a check-list manner." We support the NAS approach and see HBWoE not as an alternative to the Hill criteria, but rather as a means for rigorously and defensibly implementing them. The HBWoE emphasis on articulating the hypothesized basis and logic for why certain study results are to be taken as evidence regarding potential for human toxicity is a means for clarifying how the array of study results supports or does not support commonality and consistency across studies, while providing the logical framework for integrating specific studies within the larger body of scientific evidence to address coherence, plausibility, and analogy. HBWoE also provides a means for conducting the evaluation of each Hill criterion and – by exploring the logic of how the available data should hinge on underlying causal explanations – provides a rigorous and transparent way to express judgments about whether and to what extent the criteria are met.

In proposing his criteria, Hill emphasized that they are not independent measures to be treated as items on a checklist; rather, the criteria are guides for recognizing the possible role of hypothesized underlying causal processes as they ought to manifest themselves in the results of studies, and are to be integrated into an overall judgment based on the strength and consistency of those manifestations across the criteria. The HBWoE approach tries to provide the framework and logical basis for conducting such a critical integration, and it can be seen as a proposed means to carry out the integration of evidence across the criteria that Hill articulated.

2. Methodology

Our analysis considers the proposed hypotheses for methanol's potential carcinogenicity in humans. In evaluating a hypothesis, the following questions should be addressed:

(1) What hypothesized causal processes are seen as necessary for methanol to cause lymphoma? Are these hypothesized processes sufficient, or are other elements also necessary?

(2) For those events or processes proposed as critical, what other observable manifestations should they have (in other tissues or species)? Are these other manifestations indeed found?
(3) If either the operation or necessity of these proposed critical events were disproven, how else would one account for the array of outcomes?

The specific HBWoE evaluation approach is described in more detail in our recent publications (Rhomberg et al., 2010, 2011; Prueitt et al., 2011). To evaluate the available studies relevant to the methanol carcinogenesis, we considered the methanol carcinogenicity studies summarized in the US EPA (2009) draft document (Apaja, 1980, NEDO, 1985a,b, all cited in US EPA, 2009; and Soffritti et al., 2002). We also conducted a literature search in PubMed and Scopus using combinations of the following terms: "methanol, carcinogen*, cancer, DNA, lymphoma, metabol*, pharmacokinetic*, toxic*" (The asterisks denote terms to which truncation was applied. Variant forms of the root term were included in the search.) to identify other studies relevant to methanol carcinogenicity. Our literature searches did not identify any additional methanol carcinogenicity assays, but they did identify a key review article (Cruzan, 2009), from which additional relevant studies were identified, including genotoxicity, mode-of-action, and toxicokinetic studies. Other than the key carcinogenicity assays (Apaja, 1980, NEDO, 1985a,b, all cited in US EPA, 2009; Soffritti et al., 2002), we relied on US EPA's draft document (2009), the review article by Cruzan (2009), and all published articles since 2009 from our literature searches to obtain the most recent genotoxicity, mode-of-action, and toxicokinetic studies. All studies that directly examined genotoxicity, mode-of-action, or toxicokinetics of methanol were included in our analysis. Further, our approach does not exclude any studies a priori based on methodological or other limitations; rather, we provisionally consider the results of such studies in our subsequent evaluation of inferences, but in doing so we give them less weight by assessing their concordance with more robust studies and by recognizing that their limitations hamper their ability to discriminate between causal and artifactual explanations of their outcomes.

We reviewed and summarized the studies in the context of the proposed hypotheses regarding methanol carcinogenesis. From our review of all the relevant studies, we articulated several questions that should be considered in evaluating the evidence for carcinogenicity of methanol. We then considered the strengths and limitations of each study in an attempt to answer these questions, considering how each study informs interpretation of the other (across human, animal, exposure, and mode of action studies). We then articulated and compared alternate accounts of the existing data that have been put forth within the scientific community, considering the weight of the evidence for each account.

3. Results

3.1. Weight of evidence regarding association between exposure to methanol and carcinogenesis in humans and animals

3.1.1. Methanol carcinogenicity in humans

There are no epidemiology studies of methanol carcinogenicity in the published literature. The only available data regarding methanol toxicity in humans are from case reports and a few occupational and controlled exposure studies of neurological effects (as discussed in the recent draft toxicological review of methanol by US EPA, 2009). Because there are no human data to evaluate potential carcinogenicity from methanol exposure, the remainder of this section will focus on the data from animal studies.

3.1.2. Methanol carcinogenicity in animals

Only four carcinogenicity studies of methanol in animals have been identified, and none of these studies were peer-reviewed until recently, following considerable controversy regarding the methodology and/or results of some of these studies. The results are summarized below.

3.1.2.1. Soffritti et al. (2002) oral (drinking water) study in rats. The study by Soffritti et al. (2002) was conducted at the European Ramazzini Foundation (ERF) and was originally published as a brief report that did not undergo peer review. In this study, 100 male and 100 female Sprague–Dawley rats were exposed to 0, 500, 5000, or 20,000 ppm methanol in drinking water for 104 weeks (2 years); then, the animals were kept under observation until their spontaneous death (up to 153 weeks). Differences in tumor incidence between treated and control groups were evaluated using the χ^2 test. The authors reported increased incidences of all "hemolymphoreticular" tumors combined in high-dose females. These tumors included lymphoblastic lymphoma, lymphoblastic leukemia, lymphocytic lymphoma, lympho-immunoblastic lymphoma (LIL), histiocytic sarcoma, monocytic leukemia, and myeloid leukemia. An increased incidence of ear duct carcinomas was also reported in high-dose males.

The ERF recently made the data from the Soffritti et al. (2002) study available and the results were evaluated by Cruzan (2009). Lung pathology (including inflammation, dysplasia, or neoplasms) was present in a large majority of the treated and control rats, regardless of their age of death. For example, lung pathology was reported in 80–96% of the rats that died before 18 months and in 87–94% of all rats at study termination. A reanalysis of the tumor data by Cruzan (2009) using the standard National Toxicology Program (NTP) statistical analysis (based on Poly-3 and lifetable tests) indicated that the incidence of all hemolymphoreticular cancers combined was increased in all female treated groups compared to controls, but not in any of the treated males. Of the hemolymphoreticular tumors, 92% occurred in the lung, and the only specific cancer of this type with an increased incidence in treated animals compared to controls was LIL in high-dose males and in all treated females. The reanalysis also indicated that the incidence of ear duct carcinomas was increased in both high-dose males and females.

Cruzan (2009) noted that the methodology for carcinogenicity studies at the ERF, including the methanol study by Soffritti et al. (2002), does not conform to Organisation for Economic Co-operation and Development (OECD) (OECD, 2009) or NTP guidelines (NTP, 2010) for carcinogenicity studies. These guidelines require the termination of rat studies at 104 weeks so that the pathology of a sufficient number of controls and treated animals can be compared at the same age. The guidelines also require the sacrifice of moribund animals to aid in the minimization of tissue decay, which could interfere with pathologic evaluation.

The European Food Safety Authority (EFSA) (EFSA, 2006) proposed the hypothesis that infection in the ERF rat colony, rather than treatment, was responsible for the increased incidence in lymphomas and leukemias in the ERF studies of rats exposed to aspartame and MTBE. Others have suggested that infections in the ERF colony confounded the results of the methanol study as well (Cruzan, 2009; Schoeb et al., 2009a; Schoeb and McConnell, 2011a,b). The rat colony at ERF is not specific-pathogen-free (SPF); rather, the animals are conventionally maintained and are therefore subject to respiratory infections that can be misdiagnosed as lymphomas (Schoeb et al., 2009a; Schoeb and McConnell, 2011a).

In 2010, pathologists from NTP performed a preliminary examination of tissue slides from the Soffritti et al. (2002) study during a site visit to ERF. Histologic sections from various organs of 100 high-dose and 100 control male rats were examined and, in general, the NTP pathologists diagnosed fewer neoplasms and more

inflammatory lesions than in the original findings, noting that the inflammation was consistent with chronic infection (Malarkey and Herbert, unpublished memo). The results of this examination were not quantified in the site visit report. The NTP pathologists recommended a formal pathology peer review to resolve these discrepancies. Cruzan (2011) and Schoeb and McConnell (2011b) recently obtained, through the Freedom of Information Act (FOIA), a copy of the Slide Review Worksheet from the NTP evaluation that compared the individual slide evaluations from NTP and ERF (used in Soffritti et al., 2002). According to both evaluations, the NTP pathologists did not agree with the ERF pathologist on the majority of lymphomas and ear cancers. For example, ERF reported LIL in the lungs for 23 rats (vs. only one reported by NTP), and reported an additional 29 rats with LIL in the lungs and other tissues (vs. only 11 reported by NTP). Further, as discussed by Cruzan (2011) and Schoeb and McConnell (2011b), according to the NTP pathologists, no rats in the control group, and only one in the high-dose group (20,000 ppm methanol) had LIL in the lungs only, with five control and six high-dose rats having LIL in the lungs plus other tissues. NTP pathologists also reported six ear carcinomas in the control rats and nine in the high-dose group. Therefore, contrary to the results presented by Soffritti et al. (2002), NTP examination of the Soffritti data does not support a conclusion of increased cancer risk from methanol exposure.

3.1.2.2. Apaja (1980) dermal and oral (drinking water) study of malonaldehyde and methanol in mice. The study by Apaja (1980, as cited by US EPA, 2009) was written as a graduate thesis and the data were not published in the peer-reviewed literature. The study was not designed to evaluate the carcinogenicity of methanol; rather, the study examined the carcinogenicity of malonaldehyde (as the dimethylacetal which hydrolyzes to malonaldehyde and methanol in dilute aqueous solution) and included methanol-only controls but no concurrent negative controls. Apaja (1980, as cited by US EPA, 2009) exposed groups of 25 male and 25 female Swiss Webster mice from the conventionally maintained Eppley Institute colony to methanol in drinking water at doses of approximately 550, 1000, or 2000 mg/kg-day six times per week until their spontaneous death (up to 120 weeks). The mice in this study were euthanized when moribund. The incidence of malignant lymphomas in methanol-treated mice was higher than the overall incidence in historical controls for high-dose females and mid-dose males, but the author concluded that these incidences were "within the normal range of occurrence of malignant lymphomas in Eppley Swiss mice" (Apaja 1980, as cited by US EPA, 2009), although no data were provided to support this statement. A separate group of 25 female mice was treated by dermal exposure to approximately 530 mg/kg-day methanol three times per week until spontaneous death. The incidence of malignant lymphomas in these mice was reported to be within the historical control range, as well.

The reported incidences of lymphomas in the drinking water study were higher in methanol-only treated mice compared to those treated with both malonaldehyde and methanol, and no dose–response relationship is evident for methanol exposure and lymphomas when all groups treated with methanol (with or without malonaldehyde) are considered (see Table 1). The location of the lymphomas (in terms of organ system involved) was not reported in this study. Because the mice were not maintained under SPF conditions and there was a high level of pneumonia reported (8–28%), the Apaja (1980, as cited by US EPA, 2009) study may have involved respiratory infections similar to the ERF rat colony (TERA, 2010).

3.1.2.3. NEDO (1985a) two-year inhalation study of methanol in rats. A two-year inhalation study of methanol in rats was conducted by the New Energy Development Organization (NEDO) in Japan in the early 1980s, with the original report of the study prepared in 1985 in Japanese (NEDO, 1985a, as cited by US EPA, 2009). This study was not peer reviewed until recently (ERG, 2009), after the Methanol Institute had the study report translated into English.

Table 1
Lymphoma incidence in mice exposed to methanol and malonaldehyde in the study by Apaja (1980).[a]

Malonaldehyde dose (mg/day)	Methanol dose (mg/day)	% Lymphoma
Males		
0	82.7	17
28.0	49.8	0
0	43.5	24
17.8	31.5	12
0	24.6	4
10.3	18.2	25
Females		
0	84.5	40
23.5	41.8	4
0	40.8	36
14.8	26.2	28
0	22.6	16
8.6	15.3	36

[a] As in cited in US EPA, 2009.

In this study, groups of 52 male and 52 female F344 rats were exposed to 0, 10, 100, or 1000 ppm methanol by inhalation 19.5 h per day for 104 weeks. The study was conducted according to OECD guidelines and OECD Good Laboratory Practices (GLP). The animals were maintained under SPF conditions and were checked for antibodies to infectious agents every 4 months. The authors reported no statistically significant increases in tumor incidence in treated animals compared to controls using Fisher's exact test. As part of their recent draft toxicological review of methanol, US EPA (2009) reanalyzed the data from the 2008 translation of the study, also using Fisher's exact test, and reported increased incidences of lung adenomas and carcinomas combined in high-dose males compared to concurrent controls ($p < 0.05$). This was based on incidences of six adenomas and one carcinoma in high-dose males and of one adenoma and no carcinomas in control males. US EPA (2009) also reported an increase in the incidence of pheochromocytomas, which are tumors of the adrenal gland, in high-dose females compared to NTP historical control rates ($p < 0.05$) but not compared to concurrent controls. The authors of this study did not conduct dose–response analyses of the data.

The NEDO (1985a, as cited by US EPA, 2009) study was independently peer-reviewed in 2009 by the Eastern Research Group (ERG) at the request of US EPA (ERG, 2009). Three of five reviewers agreed with the NEDO study authors that there was no conclusive evidence of carcinogenicity. One reviewer concluded that additional analyses, such as evaluations of dose–response relationships, might have provided positive results. The final reviewer stated that induction of pheochromocytomas could not be conclusively dismissed, based on his own statistical analysis of the data which indicated a statistically significant increase of this tumor type in high-dose females compared to concurrent controls (ERG, 2009).

3.1.2.4. NEDO (1985b) 18-month inhalation study of methanol in mice. A study of chronic methanol inhalation in mice was conducted by NEDO (1985b, as cited by US EPA, 2009) during the same time period as the NEDO inhalation study in rats. Similar to the rat study, the NEDO mouse study was conducted according to OECD guidelines and OECD GLP, and was not peer-reviewed until recently (ERG, 2009).

In this study, groups of 52 male and 53 female SPF-maintained B6C3F1 mice were exposed via inhalation to the same concentrations of methanol as in the rat study (0, 10, 100, and 1000 ppm)

19.1 h per day, for 78 weeks (18 months). The authors stated that there were no increases in tumor incidence between controls and any of the treated groups. An independent peer review through ERG (2009), as well as the recent reanalysis of the data by US EPA (2009), agreed with the authors' conclusion that there were no methanol-related carcinogenic effects in mice in this study.

3.1.3. Hypothesis-Based Weight-of-Evidence evaluation of methanol carcinogenicity bioassays

Based on the available data summarized above, we ask the following questions with regard to the animal evidence for carcinogenic effects of methanol:

(1) How should the Soffritti et al. (2002) study be evaluated in light of the methodological concerns, potential respiratory infections in ERF animals, and the results of the NTP examination of the slide data from this study?
(2) How should the Apaja (1980, as cited by US EPA, 2009) study be evaluated in light of the lack of concurrent negative controls for methanol and potential respiratory infections in the animals?
(3) Should the reanalysis of the NEDO (1985a, as cited by US EPA, 2009) data in rats by US EPA (2009) be used as evidence of carcinogenicity? What is the relevance of the adrenal gland pheochromocytomas observed in this study in the context of the tumors reported by Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009)?
(4) What is the relevance of the NEDO (1985b, as cited by US EPA, 2009) mouse study?
(5) Are there alternative explanations for the increased tumor incidences (or apparent tumor incidences) in any of these studies?

Some investigators have already considered the potential problems with the chronic animal studies of methanol described above, and there is debate as to how these issues may affect conclusions regarding the potential carcinogenicity of methanol. We consider these studies and their related issues in our analysis below.

3.1.3.1. The Soffritti et al. (2002) results should not be given much weight in the evaluation of carcinogenicity. As described above, the study by Soffritti et al. (2002) was conducted at the ERF and was originally published without peer review. The study design did not conform to any established guidelines for carcinogenicity studies, and the animals were conventionally maintained. Increased incidences of all hemolymphoreticular tumors combined were observed in all treated females, but not in treated males. The vast majority of these tumors occurred in the lung, and lung pathology (particularly inflammation) was present in almost all of the rats in this study. The incidence of ear duct carcinomas was also increased in both high-dose males and females. The validity of these findings has been the subject of debate because of potential problems with the study design and the pathologic evaluation of the animal tissues.

Because the animals in the ERF rat colony are conventionally maintained rather than being kept under SPF conditions, EFSA (2006) proposed the hypothesis that infection in the colony was responsible for the increased incidence in lymphomas and leukemias in the rat carcinogenicity studies of aspartame and MTBE conducted at ERF. Others have suggested that infections, such as *Mycoplamsa pulmonis* disease, confounded the results of the ERF methanol study as well (Schoeb et al., 2009a; Schoeb and McConnell, 2011a). The hypothesis that infections were responsible for the tumor findings in ERF rat studies has sparked debate in the scientific literature. A group from US EPA (Caldwell et al., 2008) responded to this hypothesis, focusing on the studies of MTBE (Belpoggi et al., 1995, 1999).

Caldwell et al. (2008) stated that *M. pulmonis* infection was not confirmed in the rats used for the MTBE studies, but there were no reports available noting whether any testing for the infection occurred. *M. pulmonis* infection is common in conventionally-maintained rats and is the most likely cause of bronchial infection in such rats (Schoeb et al., 2009a). Approximately 77% of control and treated rats in the ERF studies of MTBE, aspartame, and methanol combined had bronchitis (Schoeb and McConnell, 2011a). Recent testimony of the lead author of the MTBE bioassay indicated that the ERF animal colony has been screened for pathogens and parasites annually since 2001, after the methanol study was completed, and antibodies to "mycoplasma" had been found (as noted by Schoeb and McConnell, 2011a).

Caldwell et al. (2008) stated that an increased incidence of hemolymphoreticular tumors is relatively rare in animals used for ERF studies and is generally associated with agents that tend to have similar chemical characteristics (such as methanol and MTBE) or metabolites in common (such as methanol, formaldehyde, aspartame, and MTBE). They noted that this apparent chemical specificity for induction of hemolymphoreticular tumor types would not be expected from a mode of action of generalized infection in the animals. In contrast, Schoeb et al. (2009a) pointed out that *M. pulmonis* infection and subsequent disease is subject to exacerbation by a variety of factors, including chemical treatment, advancing age, and husbandry conditions. Schoeb et al. (2009b) stated that it is possible that treatment with certain types of chemicals could exacerbate *M. pulmonis* disease, so that survival of ERF control rats or rats treated with other chemicals is not compromised because they aren't treated with chemicals that induce severe *M. pulmonis* disease. Furthermore, Schoeb et al. (2009b) stated that it is inappropriate to assume that chemical specificity could apply to induction of lymphomas but not to exacerbation of a respiratory infection such as *M. pulmonis* disease.

Similar to the methanol study, an increased incidence in hemolymphoreticular tumors was observed in treated female rats, but not in males, in the MTBE study. Caldwell et al. (2008) stated that if hemolymphoreticular tumors are related to a lung infection, they should occur in both sexes. Schoeb et al. (2009a) countered this argument by stating that because chemicals commonly differ in metabolism or toxicity between males and females, it is plausible that a chemical with greater toxicity in one sex could exacerbate *M. pulmonis* disease more strongly in those animals. Similar to the argument regarding chemical specificity above, Schoeb et al. (2009b) also noted that one should not assume that chemical induction of lymphomas, but not exacerbation of *M. pulmonis* disease, can differ between male and female rats. Cruzan (2009) pointed out that although the total number of hemolymphoreticular tumors was increased only in females in the methanol study, the incidence of LIL was increased in both males and females; thus, the argument by Caldwell et al. (2008) does not hold true for the methanol study.

Animals in carcinogenicity studies conducted at the ERF are kept under observation until their spontaneous death, rather than being sacrificed at 104 weeks as required by NTP (2010) and other guidelines. Caldwell et al. (2008) suggested that the increased incidences of hematolymphoreticular tumors observed in ERF studies, but not in bioassays of the same chemicals conducted at other institutions, are attributable to the longer study duration at ERF, as these tumors tend to occur more frequently in older rats. This is not supported by the data, however. Schoeb and McConnell (2011a) pointed out that in the ERF bioassays for methanol, aspartame, and MTBE, 66.5% of tumors in the lung occurred within 104 weeks. In the methanol study, there was no difference in the percentage of animals with hemolymphoreticular tumors between those dying prior to 104 weeks and those dying after 104 weeks (Cruzan, 2009). Thus, it does not appear that a study duration beyond the standard 104-week time point is advantageous for detecting tumor induction by methanol.

Caldwell et al. (2008) stated that lung infection in the ERF colony is an acute event that happens terminally in older animals and was not present throughout the animals in the MTBE study. Schoeb et al. (2009a) stated that *M. pulmonis* disease is not an opportunistic disease of aged rats but is acquired by young animals from affected mothers and persists for life. As noted above, lung pathology, including inflammation, was consistent in animals from the methanol study, regardless of age of death (Cruzan, 2009). Lung pathology was reported in 80–96% of the rats that died before 18 months and, thus, was not an old-age response (Cruzan, 2009).

The predominant form of lymphoma reported in the methanol, aspartame, and MTBE studies conducted at ERF is "lymphoimmunoblastic," which is not a recognized lymphoma type in rats; the organ most frequently affected was the lung (Cruzan, 2009; Schoeb et al., 2009a). In the methanol study, 92% of the hemolymphoreticular tumors were found in the lungs, suggesting that either the rats were highly susceptible to these tumors because of lung inflammation, or that the observations may not have been tumors (Cruzan, 2009). Other than lymphoma-like lesions reported in association with *M. pulmonis* disease, no form of lymphoma in rats has such an organ distribution (Schoeb et al., 2009a). Schoeb and McConnell (2011a) stated that the lymphoma lesions in the studies of aspartame and MTBE were reportedly characterized by accumulation of lymphocytes, plasma cells, and neutrophils, and that such lesions in the lungs of conventionally-maintained rats are more likely to be due to *M. pulmonis* disease than to chemical induction of a rare type of lymphoma with an uncharacteristic organ distribution. Overall, Schoeb et al. (2009a) and Schoeb and McConnell (2011a) consider the evidence to be convincing that the rats in the ERF bioassays of methanol, aspartame, and MTBE had *M. pulmonis* disease, and that lesions caused by this disease may have been interpreted by ERF pathologists as pulmonary lymphomas.

The results of a preliminary examination of the histologic sections from the Soffritti et al. (2002) study by pathologists from NTP provides support for the confounding of the study results by infection. As noted above, NTP pathologists diagnosed fewer leukemias and lymphomas in the lung and more inflammatory lesions consistent with chronic infection than in the original findings (Malarkey and Herbert, unpublished memo). They and others (Schoeb et al., 2009a; Schoeb and McConnell, 2011a) recommended a formal pathology peer review of respiratory sections from the studies at ERF. NTP pathologists also diagnosed "about half" of the inner ear carcinomas as reported in the original findings (Malarkey and Herbert, unpublished memo); a similar difference was found between the number of ear duct carcinomas diagnosed by a formal NTP Pathology Working Group (PWG) and ERF pathologists in the aspartame study (Cruzan, 2009). Further, as discussed above, recent review of the Slide Review Worksheets from the NTP evaluation (Cruzan, 2011; Schoeb and McConnell, 2011b) indicates that NTP pathologists did not agree with the ERF pathologist on the majority of lymphomas and ear cancers, and contrary to the results presented by Soffritti et al. (2002), the NTP examination of the Soffritti data does not support a conclusion of increased cancer risk from methanol exposure. A formal peer review of all histologic sections from the methanol study would be helpful to confirm the reported findings, as the preliminary examination by NTP was only a partial review that was "not sufficient to support or refute the overall conclusions of the studies" as reported by the ERF (Malarkey and Herbert, unpublished memo).

Another limitation of the Soffritti et al. (2002) study, and other carcinogenicity studies conducted at the ERF, is that all tumors derived from blood-forming cells are grouped together and termed "hemolymphoreticular tumors," and the incidence of this group of tumor types is reported. Individual tumor types should be considered separately, because each is derived from a different cell type and is a distinct disease with a unique mechanism of action.

Overall, the Soffritti et al. (2002) study appears to have serious methodological flaws. It was not conducted according to established guidelines for carcinogenicity bioassays, and the use of conventionally-maintained animals increases the potential for confounding of the cancer effects by infections. The likelihood of such confounding is strengthened by the presence of lung inflammation in the majority of the rats and the fact that almost all of the hemolymphoreticular cancers, including lymphomas, were found mainly in the lungs, which is uncharacteristic of lymphomas in rats. The potential misdiagnosis of lung infections as lymphomas is a serious concern, particularly in light of the diagnosis of fewer tumors and more inflammatory lesions in a subset of tissue slides from this study by NTP pathologists compared to the study authors. All of these issues make it difficult to draw conclusions about causation from this study; thus, it should not be given much weight in an evaluation of carcinogenicity of methanol.

3.1.3.2. The Apaja (1980) graduate thesis should also not be given much weight. The Apaja (1980, as cited by US EPA, 2009) study of malonaldehyde carcinogenicity in mice from the Eppley Institute has not been published in the peer-reviewed literature and was not designed to evaluate the carcinogenicity of methanol. For mice exposed to methanol in drinking water, the author reported an increased incidence of lymphomas in high-dose females and mid-dose males compared to the overall incidences in historical controls, but concluded that the incidences were still within the normal range for that strain of mice at the Eppley Institute. As noted above, no dose–response relationship was evident for methanol exposure and lymphomas in this study when all groups treated with methanol (with or without malonaldehyde) are considered.

A recent peer review of the Apaja (1980, as cited by US EPA, 2009) study, conducted through Toxicology Excellence for Risk Assessment (TERA, 2010), noted several problems with the methodology and reporting of this study, and these limit the usefulness of the findings for assessing the potential carcinogenicity of methanol. There were no concurrent negative controls in this study, so only historical control data were available for comparison. The use of historical controls likely confounds the interpretation of the data, as these controls were not matched to the experimental group in each aspect of treatment except methanol exposure. Nevertheless, the author reported that the lymphoma incidence in methanol-treated animals was still within the historical control range, although this range was not provided.

The mice in the Apaja (1980, as cited by US EPA, 2009) study were not maintained under SPF conditions and no evidence of disease surveillance in the mouse colony over the course of the study was reported. This increases the likelihood that the mice in this study may have had respiratory infections, which is further strengthened by the author's report of a high incidence of pneumonia (8–28%) in the mice. The peer review report noted that the histopathological data were not adequately described, however, and the locations of the reported lymphomas were not provided (TERA, 2010). Thus, it is not known whether any of the lymphoma lesions occurred in the lungs, or whether they were characteristic of *M. pulmonis* disease or another type of infection.

Overall, the Apaja (1980, as cited by US EPA, 2009) study does not provide evidence of a carcinogenic effect of methanol in mice. Although increased incidences of lymphomas were reported in some dose groups compared to the overall incidence in historical controls, the author stated that these incidences were within the normal range of occurrence in historical controls, and no dose–response relationship was observed. It is not known whether the use of concurrent negative controls would have provided clear evidence of carcinogenicity, or would have strengthened the findings of a lack

of carcinogenicity. It is also unclear if the findings are confounded by infections, which would strengthen the evidence for a lack of carcinogenic effects. Because of the concerns with the methodology and reporting of the Apaja (1980, as cited by US EPA, 2009) study, it should not be given much weight in the evaluation of methanol carcinogenicity. This is supported by the conclusions in the report of the peer review of this study, in which three of the four panel members stated that the study should not be used for cancer classification (TERA, 2010).

3.1.3.3. Although the NEDO (1985a) rat study should be considered as part of the weight of evidence, US EPA's reanalysis of that study should not be given much weight. Unlike the drinking water studies by Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009), the inhalation study of methanol by NEDO (1985a, as cited by US EPA, 2009) was conducted under established guidelines and the animals were maintained under SPF conditions and checked regularly for infections. The study authors reported no increased incidences of tumors, including lymphomas, in the methanol-treated rats compared to concurrent controls. The study recently underwent peer review through ERG (2009), and US EPA (2009) recently reanalyzed the data from this study after an English translation became available.

The reanalysis of data from this study by US EPA (2009) reported increased incidences of lung adenomas and carcinomas combined in high-dose males compared to concurrent controls. Only one of the lung tumors in this dose group was a carcinoma, however, and the remaining six tumors were benign adenomas. This reanalysis also indicated an increased incidence of adrenal gland pheochromocytomas in high-dose females compared to NTP historical control rates, but not compared to concurrent controls. As noted above, concurrent controls provide a better comparison, as they are matched more closely to the experimental group and are less likely to confound the interpretation of the data. Thus, the comparison to historic controls does not enhance the evaluation of carcinogenicity when data for concurrent controls are available.

In the report of the peer review of the NEDO (1985a, as cited by US EPA, 2009) study, three of the five reviewers agreed with the study authors that there was no conclusive evidence of methanol carcinogenicity, while one reviewer suggested that dose–response analyses may provide positive results, and one reviewer stated that induction of pheochromocytomas could not be conclusively dismissed after his own analysis indicated a statistically significant increase in the incidence of this tumor in high-dose females compared to concurrent controls.

The NEDO (1985a, as cited by US EPA, 2009) study should be given more weight in the evaluation of methanol carcinogenicity because of its robust study design and use of healthy animals. Overall, this study does not demonstrate a carcinogenic effect of methanol in rats. The reanalysis of the data by US EPA (2009) should not be given much weight, as it appears to have combined benign and malignant lung tumors and emphasized comparisons to historical, rather than concurrent, controls in order to report a positive result. If comparisons to historical control groups are considered, it is noteworthy that the incidence of pheochromocytomas in high-dose males was *decreased* compared to historical controls from the same studies used by US EPA (2009) for comparison to pheochromocytoma incidence in high-dose females. The US EPA (2009) reanalysis does not support carcinogenicity in the NEDO (1985a, as cited by US EPA, 2009) study, as it does not indicate an increased incidence of any malignant tumor type compared to concurrent controls. The independent statistical analysis of the pheochromocytomas by one of the peer reviewers should not be dismissed, but the statistical methods were not reported and the results are not consistent with the lack of induction of this tumor type compared to concurrent controls in the analyses by the study authors and US EPA (2009).

Moreover, in weighing the evidence for potential methanol carcinogenesis, it is not clear how the US EPA's reanalysis (if accepted), including reported observations of one lung carcinoma and adrenal gland pheochromocytomas in high-dose females, fits with the Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009) results in which reported tumors were lymphomas and other hemolymphoreticular tumor types. The lack of consistency regarding reported tumor types across these studies weakens the overall weight of evidence for methanol carcinogenesis. That is, why should one assume the reported lymphomas in the Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009) studies are the tumor type of concern, when they carry no more weight (and possibly less, given the methodological flaws and potential confounding effects is these studies) than the reanalysis of the NEDO (1985a) study by US EPA where lung adenomas and adrenal gland pheochromocytomas have been reported? Therefore, in addition to potential problems with each individual study or reanalysis, the inconsistencies across the studies further weakens the evidence for methanol carcinogenesis.

3.1.3.4. The NEDO (1985b) mouse study results need to be considered in the weight of evidence. Similar to the NEDO (1985a, as cited by US EPA, 2009) rat inhalation study, the NEDO (1985b, as cited by US EPA, 2009) mouse inhalation study was robust and the animals were SPF and routinely checked for infections. The authors reported no increased incidences of tumors in any of the methanol-treated groups compared to the concurrent controls. These conclusions were consistent with those of an independent peer review of the study through ERG (2009) and of a recent review of the data by US EPA (2009).

The NEDO (1985b, as cited by US EPA, 2009) mouse study should be given more weight in the evaluation of methanol carcinogenicity because of the robust study design and use of healthy animals. This study provides evidence that there is no carcinogenic effect of methanol in mice.

3.1.4. Summary and conclusions

There are no epidemiology studies of methanol carcinogenicity available, so the weight of evidence regarding potential methanol carcinogenicity in whole organisms must rely on the available animal data. These data are limited, however, as only four carcinogenicity studies in rodents have been identified: chronic drinking water studies in rats (Soffritti et al., 2002) and mice (Apaja, 1980, as cited by US EPA, 2009), and chronic inhalation studies in rats (NEDO, 1985a, as cited by US EPA, 2009) and mice (NEDO, 1985b, as cited by US EPA, 2009).

The Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009) drinking water studies have serious methodological flaws, and the potential confounding effects from suspected respiratory infections make it difficult to draw any conclusions about causation from these studies. Thus, although they need to be considered, these studies should not be given much weight in the overall weight-of-evidence evaluation of methanol carcinogenicity. Further, even if the methodological issues in the Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009) studies are overlooked, the data from these studies do not fit with other evidence from studies in healthy animals. The induction of hemolymphoreticular and ear duct tumors in rats in the Soffritti et al. (2002) study and of lymphomas in the Apaja (1980, as cited by US EPA, 2009) study are not consistent with the NEDO (1985a, as cited by US EPA, 2009) rat study and NEDO (1985b, as cited by US EPA, 2009) mouse study results.

The NEDO (1985a,b both as cited by US EPA, 2009) rat and mouse inhalation studies should be given more weight in the overall weight-of-evidence evaluation because they are robust and there was no evidence of disease in the animals. These studies

reported no carcinogenic effects, of any of the tumor types reported by Soffritti et al. (2002) or Apaja (1980, as cited by US EPA, 2009), from chronic exposure to methanol. Although US EPA (2009) reanalyzed the NEDO (1985a, as cited by US EPA, 2009) rat study, the reanalysis should not be given much weight because it appears to have combined benign and malignant lung tumors and emphasized comparisons to historical, rather than concurrent, controls in order to report a positive result. The increased incidence of pheochromocytomas compared to historical controls is the only potential carcinogenic effect from the US EPA (2009) reanalysis of this study, and adrenal gland pheochromocytomas are very different from the lymphomas and other hemolymphoreticular tumor types reported by Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009), so it is not clear how these data support each other in the overall weight of evidence.

Without giving much weight to the Soffritti et al. (2002) study and the Apaja (1980, as cited by US EPA, 2009) thesis, we are left with the NEDO (1985a,b both as cited by US EPA, 2009) studies in rats and mice, which do not report a statistically significant carcinogenic effect of methanol. And even if we accept the US EPA (2009) reanalysis of the NEDO (1985a, as cited by US EPA, 2009) rat study, we are left with one positive study (adrenal gland pheochromocytomas in rats) and one negative study (in mice). Therefore, the more robust studies (the NEDO studies) suggest either no carcinogenic effect (based on the original reported results), or inconsistent effects (if one accepts US EPA's reanalysis of the rat data), at best, from chronic exposure to methanol. Further, even the inconsistent effects within the NEDO studies are not consistent with the effects in the Soffritti et al. (2002) study or the Apaja (1980, as cited by US EPA, 2009) thesis.

Although it is reasonable to consider that the differences observed between the drinking water and inhalation studies may be related to the different routes of exposure, as discussed in the next section, once methanol is absorbed (via inhalation or oral exposure), it is rapidly distributed to all organs, and the tissue:blood concentration ratios are predicted to be similar. Therefore, similar effects from both routes of exposure are expected. We recommend a formal pathology peer review of the Soffritti et al. (2002) data and additional methanol bioassays in healthy animals (via drinking water), in accordance with OECD or NTP guidelines (depending on sponsor). This information will be helpful to confirm the results of the NEDO (1985a,b, both as cited by US EPA, 2009) rat and mouse studies and to confirm that different routes of exposure result in similar effects.

Further, it is important to evaluate the reported carcinogenic effects in the context of whether there is a plausible biological mode of action for methanol carcinogenesis. That is, if there is a plausible mode of action, then one might give more weight to the carcinogenesis data in the overall weight of evidence. As discussed in the next section, there is little support in the scientific literature for the proposed modes of action for methanol carcinogenesis. It is also important to consider that methanol does not have the typical properties of chemicals that, upon exposure, cause lymphohematopoietic cancers (such as hematotoxicity or bone marrow toxicity). Neither of these effects was observed in the animal bioassays described herein. Therefore, the overall weight of evidence, including the lack of evidence for a plausible mode of action, does not support a causal association between chronic methanol exposure and carcinogenesis.

3.2. Weight of evidence regarding a plausible mode of action for methanol carcinogenesis

As discussed in the previous section, the overall weight of evidence from the methanol toxicological data, and the lack of human data, provides little support for a causal association between exposure to methanol and carcinogenesis (particularly lymphoma) in animals and humans.

Despite these findings, two potential modes of action for methanol-induced lymphoma have been put forth:

(1) Methanol metabolism to formaldehyde, followed by formaldehyde induction of lymphoma via one of the modes of action recently proposed in US EPA's draft toxicological review of formaldehyde (US EPA, 2010).
(2) Induction of carcinogenesis in rodents via oxidative stress caused by hydrogen peroxide release during catalase-induced metabolism of methanol to formaldehyde.

To evaluate the weight of evidence with regard to the plausibility of these proposed modes of action, below we summarize what is known about toxicokinetics and DNA reactivity of methanol in animals and humans, followed by what is known about the lymphohematopoietic carcinogenic potential of formaldehyde in animals and humans. In light of what it is known, and what has been proposed, we set out several questions that should be considered in evaluating the weight of evidence for the plausibility of these two modes of action.

3.2.1. Methanol toxicokinetics

The toxicokinetics of methanol has been extensively studied and are summarized in a recent review by Cruzan (2009) and in US EPA's draft toxicological review (2009). Methanol is endogenously produced and therefore naturally present in animals and humans. It is also naturally present in the human body due to dietary sources such as fruits, vegetables, and alcoholic beverages. Methanol is needed for "one-carbon" additions to nucleic acids, proteins, etc. Once methanol is absorbed (via inhalation or oral exposure), it is rapidly distributed to all organs, and the tissue:blood concentration ratios are predicted to be similar via different routes of exposure. The human blood level of methanol typically ranges from 1 to 4 mg/L (Cruzan, 2009). Under physiological conditions, methanol is metabolized to formaldehyde, which is then oxidized to formic acid, and further to carbon dioxide. The metabolism of methanol in rats and mice is different from that in humans. In rats and mice, methanol can be metabolized to formaldehyde via catalase or alcohol dehydrogenase (ADH1), with metabolism by catalase resulting in hydrogen peroxide as a byproduct. In humans, only ADH1 is involved in the metabolism of methanol to formaldehyde. Once formaldehyde is formed, formaldehyde dehydrogenase and *S*-formylglutathione hydrolase are involved in rapidly metabolizing formaldehyde to formic acid in both animals and humans. Formic acid is then converted to carbon dioxide via a folate-dependent enzyme system pathway in humans and a folate-dependent enzyme system and CAT-peroxide system in animals (Cruzan, 2009; Hayes, 2001; US EPA, 2009).

The half-life of administered methanol (approximately 3 h) is much longer than the half-life of formaldehyde (approximately 1 min) (Cruzan, 2009; US EPA, 2009); thus, accumulation of formaldehyde in tissues, as a result of typical methanol exposures, is not likely. Typical environmental exposures to methanol for humans are likely less than 22 mg/kg-day, based on the OSHA methanol standard of 200 ppm,[1] and more typically are in the range of 0.3–1.1 mg/kg-day from dietary intake of aspartame (Kavet and Nauss, 1990). The typical "background" body burden of methanol is 0.5 mg/kg (Kavet and Nauss, 1990).

[1] The OSHA standard for methanol is 200 ppm (NIOSH, 2007), which equals 262 mg/m^3. If we assume an inhalation rate of 10 m^3/day for workers, a body weight of 70 kg, and a 60% absorption rate from inhalation (Kavet and Nauss, 1990), this results in a "worst-case" dose of 22 mg/kg-day methanol.

3.2.2. Methanol genotoxicity

In rats and mice, catalase metabolism of methanol to formaldehyde produces hydrogen peroxide as a byproduct (Kavet and Nauss, 1990). Since hydrogen peroxide (a reactive oxygen species) induces oxidative stress, often leading to lipid peroxidation and oxidative DNA damage in cells where it is produced, oxidative stress has been proposed as a potential mechanism of action for methanol carcinogenesis (US EPA, 2009). Parthasarathy et al. (2006) observed increased lipid peroxidation products in lymphoid tissue in rats exposed to 2370 mg/kg-day methanol for 15 and 30 days, but more recent studies show that treatment of mice, rabbits, and monkeys with similar doses of methanol for 15 days (2 g/kg i.p.) (McCallum et al., 2011a,b) did not result in increased levels of oxidative DNA damage (as measured by 8-oxo-dG) in bone marrow, spleen, kidney, lung, or liver in any species tested. In addition, as part of one of the studies (McCallum et al., 2011a), fibroblasts from DNA-repair deficient mice [oxoguanine glycosylase (*Ogg1*) knockout (KO) mice deficient in the enzyme responsible for the repair of 8-oxo-dG DNA damage] were treated with methanol and did not accumulate 8-oxo-dG DNA damage. Further, in vivo exposure of *Ogg1* KO mice to methanol (2 g/kg i.p.) did not result in increased levels of 8-oxo-dG in bone marrow, spleen, kidney, lung, or liver (McCallum et al., 2011a,b). Overall, these data suggest that methanol does not induce oxidative DNA damage.

Methanol genotoxicity assays have been predominantly negative, as discussed in Cruzan (2009) and in US EPA's draft toxicological review (US EPA, 2009). In vitro assays in bacteria and mammalian cell culture [Chinese hamster lung (CHL) cells, mouse lymphoma cells, Syrian hamster embryo cells, and human cells] are more than 75% negative, including negative results in human cell culture. The NEDO (1987, as cited in Cruzan, 2009 and US EPA, 2009) genotoxicity assays in CHL cells were negative for forward mutations and chromosomal aberrations (CA), but did display a slight increase in sister chromatid exchanges (SCE) in CHL cells at the highest methanol concentration (28.5 mg/L).

In vivo methanol genotoxicity studies have been almost entirely negative (in mice), as discussed by Cruzan (2009) and US EPA (2009), and in the recent McCallum et al. (2011a,b) studies, except for two old abstracts (Chang et al., 1983; Pereira et al., 1982) whose authors claimed to have observed increased CA in animals from methanol exposure, but not an increase in micronuclei (MN); these studies have not been found in published literature. Wang et al. (2008) measured formaldehyde-induced DNA adducts [deoxyadenosine (dA) adducts in rats after formaldehyde or methanol exposure (via drinking water) and found no increase in these adducts in either leukocytes or liver from methanol or formaldehyde exposure. A recent study by Gul et al. (2011), however, shows relatively small increases above endogenous levels of formaldehyde-induced dG adducts in rats exposed to high levels of methanol (500–2000 mg/kg-day; much higher than typical human exposures which is closer to 22 mg/kg-day, as discussed above), but, even at these high doses, the endogenous adducts still outnumber the exogenous adducts. NEDO (1987, as cited in Cruzan, 2009 and US EPA, 2009) found no increase in MN formation in SPF mouse bone marrow cells upon gavage dosing with concentrations of methanol (1050, 2110, 4210, and 8410 mg/kg) comparable to (or higher than) the doses used by Gul et al. (2011).

3.2.3. Lymphohematopoietic carcinogenic potential of formaldehyde

As discussed by US EPA (2009) and Cruzan (2009), one proposed mechanism is that formaldehyde, generated from the metabolism of methanol, is the ultimate carcinogen from methanol exposure. Therefore, it is important to consider the weight of the available evidence put forth as supportive of a causal association between formaldehyde exposure and lymphoma, particularly hemolymphoreticular cancers of the lung that have been suggested to be causally associated with methanol exposure. Below, we summarize what is known about the potential for formaldehyde to cause lymphoma.

3.2.3.1. The weight of evidence does not support a causal association between formaldehyde exposure and lymphoma. As discussed in several recent review articles (Cruzan, 2009; Golden et al., 2006; Pyatt et al., 2008; Rhomberg et al., 2011), the weight of evidence does not support a causal association between formaldehyde exposure and hemolymphoreticular cancers (particularly lymphomas). In fact, the current debate regarding formaldehyde lymphohematopoietic carcinogenesis is whether formaldehyde exposure causes increased risk of leukemia, not lymphoma. Since leukemias and lymphomas derive from different cells types, they should be considered separately (as discussed above), and consequently, the formaldehyde human data do not support a carcinogenic mechanism for methanol and lymphomas.

Cancer-type aside (i.e., lymphoma or leukemia), the weight of evidence does not support a causal association between formaldehyde exposure and any type of lymphohematopoietic cancers. As summarized in our Hypothesis-Based Weight-of-Evidence (HBWoE) evaluation of formaldehyde leukemogenesis (Rhomberg et al., 2011), there is no lymphohematopoietic cancer or group of lymphohematopoietic cancers for which associations with formaldehyde were found consistently within or across human epidemiology studies. Although some statistically significant associations were reported, these were outnumbered by null findings using related exposure metrics, and there were no consistent exposure–response relationships observed. Because of the lack of consistency across studies, it is most likely that any observed effects were a result of confounders, limitations in statistical methods (e.g., multiple comparisons), disease misclassification, and/or exposure misclassification/measurement error. Furthermore, the available studies of formaldehyde hematotoxicity in both animals and humans provide little evidence for formaldehyde-associated leukemia, with animal studies generally reporting neither hematotoxicity nor leukemia associated with formaldehyde inhalation or ingestion. And finally, based on evidence suggesting that formaldehyde does not move beyond the nasal respiratory mucosa to increase levels in the blood, and the lack of evidence for distant site formaldehyde-induced DNA damage beyond the nasal mucosa, the overall weight of evidence provides little support for the three proposed modes of action for formaldehyde leukemogenesis: (1) directly targeting bone marrow hematopoietic stem cells; (2) targeting nasal stem cells, in nasal-associated lymphoid tissue (NALT), that later home to bone marrow; or (3) targeting circulating hematopoietic stem cells that later home to bone marrow.

3.2.4. Hypothesis-Based Weight-of-Evidence evaluation of the proposed modes of action for methanol carcinogenesis

The plausibility of the two proposed modes of action for methanol carcinogenesis have been extensively reviewed by Cruzan (2009), and also as part of comments provided by the Methanol Institute (MI) (MI, 2010) on US EPA's Draft Toxicological Review of Methanol (US EPA, 2009). We have considered the data summarized above and have come to the following questions with regard to the proposed modes of action.

With regard to the proposed mode of action involving formaldehyde as the ultimate carcinogen, several key questions need to be considered:

(1) How does one account for the different tumor types reported from formaldehyde exposure (predominantly leukemias) vs. methanol exposure (lymphomas)?
(2) Is it biologically plausible for methanol exposure, at typical environmental exposure levels, to lead to concentrations of formaldehyde in tissues above endogenous levels?

(3) How does one account for differences in genotoxic effects between methanol and formaldehyde, if formaldehyde is ultimately leading to the mutations proposed to induce carcinogenesis from methanol exposure?
(4) How does one account for the weight of evidence supporting a lack of a causal association between formaldehyde and lymphoma, or any lymphohematopoietic cancer?

With regard to the proposed mode of action involving oxidative stress from catalase production of hydrogen peroxide, additional key questions need to be considered:

(1) What is the evidence that methanol exposure leads to conditions of oxidative stress?
(2) Since this mechanism is based on induction of oxidative stress from hydrogen peroxide released during catalase metabolism of methanol to formaldehyde, and humans utilize only ADH1 to metabolize methanol to formaldehyde, what is the human relevance of this proposed mechanism?

3.2.4.1. There is no evidence to support a mode of action for methanol carcinogenesis that involves formaldehyde as the ultimate carcinogen. The weight of evidence does not support a mode of action for methanol carcinogenesis involving formaldehyde as the ultimate carcinogen. First, the tumor types reported in the methanol toxicity studies are lymphomas, which are different from leukemias predominantly reported in the formaldehyde epidemiology and toxicity studies, and as stated above, these two types of lymphohematopoietic tumors derive from different cell types and need to be considered separately. Therefore, the debate regarding formaldehyde leukemogenesis aside, the toxicological and epidemiological formaldehyde carcinogenesis data (if accepted) do not support an association between methanol exposure (via metabolism to formaldehyde) and increased risk of lymphoma. Further, of lymphohematopoietic cancers, leukemias are primarily caused by chemical agents and lymphomas are more likely to be associated with infectious agents (viruses or bacteria) (US EPA, 2010); as discussed above, it is possible that the animals in the Soffritti et al. (2002) and Apaja (1980, as cited by US EPA, 2009) studies had respiratory infections.

Second, if methanol carcinogenesis involves formaldehyde as the ultimate carcinogen, one would expect similar genotoxic effects for methanol exposure and formaldehyde exposure, and this is not the case. High exposures to formaldehyde have been shown to induce DNA–protein crosslinks and formaldehyde–DNA adducts in the nasal mucosa of rats, and in bacterial and mammalian in vitro assays (Rhomberg et al., 2011; US EPA, 2010). As already discussed, however, methanol genotoxicity assays have been essentially negative and, although high exposure concentrations may lead to a small increase in formaldehyde DNA adducts (Gul et al., 2011), higher exposure concentrations were not found to be genotoxic or tumorigenic (NEDO, 1987, as cited in Cruzan, 2009 and US EPA, 2009). Further, no studies have reported formaldehyde-induced DNA–protein crosslinks from methanol exposure.

Third, central to the plausibility of the formaldehyde mode of action is that methanol exposure would necessarily need to lead to levels of formaldehyde in the blood above endogenous levels. As discussed above, given that methanol is metabolized slowly to formaldehyde (half-life of 3 h) and then formaldehyde is quickly metabolized to formic acid (half-life of 1 min), it is not biologically plausible that formaldehyde, generated as the result of typical human methanol exposure levels, will accumulate in tissues above endogenous levels. In fact, there is no evidence presented by US EPA (2009) or others suggesting that typical human methanol exposures lead to increased levels of formaldehyde in the blood beyond endogenous levels. There is, however, a lack of evidence supporting increased levels of formaldehyde in the blood even from direct exposures to high concentrations of formaldehyde. Recent studies in rats and monkeys by Lu et al. (2010, 2011) and Moeller et al. (2011), using deuterium-labeled formaldehyde (i.e., [$^{13}CD_2$]formaldehyde), indicate that inhalation exposure to high concentrations of formaldehyde (0.7–15.2 ppm, 6 h/day, from 1 to 5 days) does not cause DNA damage beyond the portal of entry; i.e., exogenous formaldehyde-induced DNA adducts were not observed in the bone marrow, circulating hematopoietic stem cells, or the NALT. Although one could argue that methanol travels to different target tissues where there is then local metabolism to formaldehyde and subsequent DNA damage by formaldehyde, as discussed, the evidence does not suggest that blood or tissue levels of formaldehyde are increased above endogenous levels following methanol exposure (US EPA, 2009), nor does it suggest that formaldehyde DNA adducts are increased above endogenous levels following high exposures to methanol (Gul et al., 2011). These studies provide strong evidence to support the biological implausibility of inhaled methanol, at typical human exposure levels, to result in increased blood and tissue levels of formaldehyde and distant site carcinogenicity, such as leukemia or lymphoma.

As discussed above, Wang et al. (2008) measured formaldehyde-induced dA adducts in rats after formaldehyde or methanol exposure and found no increase in adducts in either leukocytes or liver from methanol or formaldehyde. As already discussed, a recent study by Gul et al. (2011) found relatively small increases above endogenous levels of formaldehyde-induced deoxyguanosine (dG) adducts in rats exposed to high concentrations of methanol. Even at these high doses, however, the endogenous adducts still outnumbered the exogenous adducts. Although these data suggest the possibility that, at least in rats, exposure to very high concentrations of methanol could lead to small increases above endogenous levels of formaldehyde in the blood and in the levels of formaldehyde DNA-adducts, the data do not support a high mutagenic/genotoxic potential for this level of adducts, consistent with the lack of genotoxicity observed in the majority of methanol genotoxicity studies. For example, NEDO (1987, as cited in Cruzan, 2009 and US EPA, 2009) found no increase in MN formation in SPF mouse bone marrow cells upon gavage dosing with concentrations of methanol (1050, 2110, 4210, and 8410 mg/kg) comparable to (or higher than) the doses used by Gul et al. (2011); the NEDO study also found no increase in tumorigenesis. Therefore, much lower, typical human exposures to methanol (less than 22 mg/kg-day) would likely be insufficient to generate biologically relevant levels of formaldehyde DNA adducts (i.e., the level of exogenous adducts would be extremely small compared to the level of endogenous adducts, as suggested by Gul et al. (2011), and consequently, would not induce sufficient genotoxicity above background levels to cause an increased risk in tumorigenesis.

Further, although studies suggest that the metabolism of methanol is similar in rodents and non-human primates, even though the first step in the metabolism occurs through different pathways (catalase in rodents vs. ADH1 in humans and non-human primates) (Kavet and Nauss, 1990), it is worth considering that at very high exposure concentrations, where metabolic saturation begins to occur, resulting in increased blood levels of methanol, that perhaps there are differences between species that could result in different rates of methanol metabolism to formaldehyde. In fact, physiologically based pharmacokinetic (PBPK) models for methanol in rats, monkeys, and humans (Bouchard et al., 2001; Horton et al., 1992) show that metabolism is very similar across species at concentrations less than 1200 ppm, but blood methanol levels begin to differ across species at higher exposure concentrations. Therefore, the potential for (or lack of potential for) species differences in methanol metabolism should be considered as part of interpretation of results from very high methanol exposure concentrations.

Finally, and also central to the lack of plausibility of this mode of action, is that the scientific basis for an association between formaldehyde exposure and lymphohematopoietic carcinogenesis (predominantly leukemogenesis) is weak. Although there is considerable debate regarding the possibility of an association, the overall weight of evidence provides little support for formaldehyde leukemogenesis (Rhomberg et al., 2011). Therefore, until the weight of evidence shifts more in favor of an association between formaldehyde exposure and increased risk of lymphoma (and currently this is unlikely since there is only weak to no evidence for an increased risk of a different type of lymphohematopoietic cancer from formaldehyde exposure – leukemia), there is little scientific basis for entertaining the possibility that methanol metabolism to formaldehyde leads to an increased risk of lymphoma.

3.2.4.2. There is no evidence to support a mode of action for methanol carcinogenesis that involves oxidative stress, particularly in humans. There is little support for a proposed mode of action in animals (and no support in humans) for methanol carcinogenesis involving the production of reactive oxygen species and oxidative stress from the metabolism of methanol to formaldehyde via catalase which produces hydrogen peroxide as a byproduct.

As discussed above, the animal data do not support an association between methanol exposure and oxidative stress. Although some evidence of oxidative stress has been observed in rats exposed to high concentrations of methanol (Parthasarathy et al., 2006), exposure to similar levels in rabbits, mice, and monkeys did not increase oxidative DNA damage (as measured by 8-oxo-dG adducts) in lung, liver, or bone marrow cells (McCallum et al., 2011a,b).

The lack of evidence for an oxidative stress mechanism for methanol carcinogenesis in animals aside, this proposed mechanism is not relevant in humans since ADH1, and not catalase, is involved in the metabolism of methanol to formaldehyde (Hayes, 2001; US EPA, 2009), and therefore hydrogen peroxide is not a byproduct of methanol metabolism in humans. Consequently, the weight of evidence strongly contradicts an oxidative stress mode of action for methanol carcinogenesis in humans.

3.2.5. Summary and conclusions

The proposed modes of action for methanol carcinogenesis (lymphohematopoietic cancer) find little support in the scientific literature.

First, with regard to the proposed mode of action involving formaldehyde, there is no evidence that typical environmental exposures to methanol lead to levels of formaldehyde in blood or tissues that are above endogenous levels; in fact, the evidence suggests this is biologically implausible. Without increased levels of formaldehyde, the proposed mode of action involving formaldehyde as the ultimate carcinogen is not supported. Further, the biological implausibility of the involvement of formaldehyde in methanol carcinogenesis is supported by toxicological observations. That is, the tumor types reported from methanol exposure (lymphomas) are different from the tumor types reported from formaldehyde exposure (leukemias), and the two tumor types should be considered separately. And finally, the current weight of evidence provides little support for an association between formaldehyde exposure and increased risk of any lymphohematopoietic cancer (leukemia or lymphoma). Therefore, with little basis for asserting that formaldehyde exposure leads to leukemia or lymphoma, there is little, if any, scientific basis for asserting that the generation of formaldehyde from methanol metabolism is causally associated with the reported increase in methanol-induced lymphomas.

In addition to the lack of evidence supporting a biologically plausible mode of action, as discussed earlier, methanol does not have the typical properties of chemicals that, upon exposure, cause lymphohematopoietic cancers (such as hematotoxicity or bone marrow toxicity). Neither of these effects was observed in the animal bioassays described herein.

Further, animal data provide little support for the proposed mode of action involving oxidative stress from catalase production of hydrogen peroxide during metabolism of methanol to formaldehyde. Moreover, given that humans do not utilize catalase for this metabolism (i.e., methanol is metabolized to formaldehyde via ADH1), the weight of evidence strongly contradicts an oxidative stress mode of action in humans.

Finally, the lack of evidence for the two proposed modes of action for methanol carcinogenesis is supported by weak evidence of methanol carcinogenesis in animals (as discussed earlier), where the only positive studies are possibly confounded by chronic respiratory infection in the animals. If the carcinogenesis data were more convincing, they would provide more of a basis for proposing a carcinogenic mode of action. Taken as a whole, the inconclusive evidence of carcinogenesis in animals, together with lack of evidence for methanol-induced genotoxicity, and evidence that the proposed modes of action are either biologically implausible or lack human relevance, suggests that the overall weight of evidence for a causal association between methanol exposure and increased risk of lymphoma (particularly in humans) is not supported.

4. Discussion

HBWoE comes down to evaluation of alternative "accounts" (Rhomberg et al., 2010, 2011; Prueitt et al., 2011). An account (which we put forth in this context as a technical term), is a proposed set of explanations for the set of observed phenomena across the body of relevant observations. The explanations need not be proven – what is important is that one set out what is being proposed as causal and generalizable phenomena (that constitute the basis for applying observations of biological perturbations or realized risks in other contexts to project potential risks to humans as they are exposed), what is being proposed as the basis for deviations that lead to observations that do not fit the hypothesized causal model (i.e., that would otherwise be counterexamples or refutations), what assumptions are made that are ad hoc (to explain particulars, but for which the evidence consists of their plausibility and the observations they are adduced to explain), what further auxiliary assumptions have to be made (and how reasonable they are in view of our wider knowledge and understanding), and what is relegated to error, happenstance, or other causes not relevant to the question at hand. There are competing accounts, and one should evaluate the main accounts as to: how the evidence supports them; what is necessary to assume; and overall, how the weight of evidence for each suggests how compelling the account is.

The importance of considering alternative "accounts" is made explicit by Austin Bradford Hill in his seminal paper on distinguishing causality from association:

> *None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non. What they can do, with greater or less strength, is to help us to make up our minds on* ***the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?*** (Hill, 1965) [emphasis added].

The essence of the accounts is that they constitute being explicit about Hill's "ways of explaining the set of facts before us." They are not conclusions or findings, but rather provisional proposals for the reasons behind the set of observations at hand – noting proposed

causal pathway commonalities among different observations and proposing reasons, even if only tentative and presumed, for why apparent contradictions to the general operation of those causal principles are to be provisionally accounted for, set out in a way that makes clear where assumptions, interpretations, and tentative inferences have been drawn. It is by comparing alternative accounts – alternative hypotheses about what causal effects actually exist – and assessing their comparative support from information at hand that we can judge how compelling each alternative should be deemed to be and, hence, with what degree of confidence we can judge the hypothesized causal processes (and their consequences for human risk estimation) to be supported by the factual record. The support for each account is to be assessed by examining its success at explaining phenomena, its need for assumptions to fill in gaps (and the comparative reasonableness of those assumptions, compared to those needed for competing accounts), and its comparative invocation of ad hoc suppositions that are necessary to accommodate what might otherwise be inexplicable results. An account is most compelling when it not only stands in accord with the data, but also helps "explain" the data in a parsimonious way (not elaborating a separate reason for every observation but instead finding common reasons for sets of observations) and, moreover, achieves this explanatory ability much more readily than any competing account. The alternative accounts reflect the set of hypotheses and what reasons one in effect accepts as true for each set of available associated observations if one subscribes to the account.

4.1. Consideration of methodological limitations in critical studies in considering alternative accounts

Methanol presents an interesting example of a more general issue in carcinogen hazard identification: the problem of how to consider the impact of studies with considerable methodological limitations. One approach is to conduct an initial quality screen, and eliminate studies from consideration that fail to meet set standards of conduct. The difficulty with this approach is that it is hard to completely reject the possibility that the omitted studies provide at least some information about real phenomena despite their failings, and this is exacerbated when (as with methanol) there are few rigorously conducted alternative studies on which to rely. The approach risks criticism that inconvenient results at odds with a predetermined conclusion are being disqualified on technicalities. Another approach is to include the questioned studies, but somehow to discount their weight in the overall conclusion. When, as with methanol, the collective evidence reveals only questioned suggestions of effects of potential relevance, one can conclude that no rigorous basis for identifying a human hazard potential exists and assign a low likelihood to the agent's human carcinogenicity. The difficulty with this approach is that, without an overt process for conducting the weighing, the conclusion is also subject to criticism on the grounds that studies were either over- or underweighed, and it is not clear what to make of down-weighted but nonetheless non-rejected possibilities, or why the overall conclusion is the best interpretation of the data at hand.

In our view, our hypothesis-based approach to weight of evidence provides an improved approach; it neither eliminates studies outright nor takes them on face value. By making the inference process explicit, and by calling attention to the reasoning behind a weight-of-evidence conclusion and specifically how it is connected to study results, it makes the process transparent and subject to legitimate scientific discussion about how compellingly data support specific interpretations. The key to the process is to be explicit about the reasoning behind why particular study outcomes are to be seen as evidence of a potential carcinogenic effect in humans, and to recognize such inference as a generalization (and not just an extrapolation) – that is, as a hypothesized feature of the interaction of the agent with mammalian biology so as to lead to risk of generating tumors. As a generalization, this feature ought to apply to all tested animals, or if it does not, there should be biological reasons behind the exceptions, and there should be other observable consequences of the hypothesized processes, evidence for or against which becomes part of the weight-of-evidence consideration. In the hypothesis-based approach, parallelism or congruence of effects (either frank endpoints or evidence of the operation of the hypothesized causative processes) among datasets strengthens the argument for the existence and operation of the hypothesized risk-generating processes, and lack of concordance tends to refute its general existence, and hence the strength of its presumed application to humans. The reasoning as to why incongruities and exceptions exist, at least in the form of hypothesized bases for their failure to conform to the hypothesized generalization, become part of the weight of evidence.

Each study outcome of potential relevance has alternative interpretations that must be weighed against one another – that the effect is real and is a manifestation of some process that the test animals have in common with humans, suggesting that humans might share the chemically-induced effects, or that the effect is an artifact of the study shortcomings and does not represent an instance of a tumor-generating process even in the study animals themselves, much less in humans, or even that the effect in the test animals is real enough, but its mode of action is such that it is a manifestation of a process specific to that strain or set of conditions and would not constitute a process generalizable to other systems, and therefore does not constitute a basis for inference of human cancer risk. The patterns of congruence and mutual support – or the lack thereof – among all the studies in the body of evidence, provides information about the relative plausibility of these alternative explanations of individual results, and the weight of evidence as a whole comprises a comparative analysis of the relative credence one should place in the alternative accounts of the whole array of study outcomes at hand (in view of all of the data and the collections of explanations of the observed outcomes that constitute alternative accounts).

In this view, the way to consider studies with methodological problems is to consider the alternative hypotheses that: (a) despite the problems, the study results indicate a particular effect of human relevance, in view of the database-wide evidence for and against that effect, and the potential methodological problems have not effaced or biased the particular result; or (b) the study's apparent effect is in fact an artifact of its methodological limitations, in view of its incongruence with other study outcomes and the perceived likelihood that the methodological flaws could produce such an artifact. This alternative explanation is strengthened to the extent that one can adduce evidence for the actual operation (and not just hypothetical role) of the factors being deemed to have resulted in causing the artifactual results, for instance evidence that other expected consequences of the proposed artifact-generating processes are observed. In short, rejection of study results is itself a proffered hypothesis to be evaluated against available data. What makes a poor study poor, or an unreliable study unreliable, is that the design or conduct flaws lead to a hampered ability to ascribe results to the causes one intends to test vis-à-vis to artifactual causes. Good study design and conduct strive to reduce this ambiguity. That is, in HBWoE, poor studies are automatically downweighted by their inability to differentially add weight to the causative versus noncausative accounts.

In the case of methanol's weight of evidence, it is our view that the lack of consistent connection or corroboration among the study results, combined with the ready and plausible alternative explanations that the apparent results are artifacts of poor experimental design and conduct – and further combined with the mechanistic

implausibility of the apparent results in view of known biology and properties of methanol – that lead us to conclude that no compelling animal evidence of methanol's human carcinogenicity exists. That is, we deal with the imperfect studies by explicitly considering and preferring the alternative explanation of apparent positive results as artifacts, finding this explanation more compelling than that they are true human-relevant effects that have shown through despite the study shortcomings, and that the difficulty in imagining a mechanistic basis for the apparent actions can be overcome by proposing that we are somehow mistaken about the severity of the difficulties. In short, the interpretation of apparent positive animal results as artifacts is judged more plausible than that the disparate effects are real and indicators of potential risks in humans, with the differences among studies and the mechanistic difficulties explicable by an extensive set of special pleadings as to why they should not be refuting.

4.2. Comparison of alternative accounts for reported effects from methanol exposure

For methanol, there are two accounts that need consideration. One account consists of acceptance of the animal evidence as sufficiently compelling that, even in the face of no human data, with the concerns regarding the potential flaws and confounding effects in the animal studies, and inconsistency of effects across studies (if one accepts the reported effects), one of the proposed modes of action for methanol carcinogenesis must be right. Moreover, the arguments against the biological plausibility of these modes of action must in some way be incorrect. Acceptance of this account is associated with many unanswered questions and post hoc explanations for how the current data should be interpreted as supporting it. This account requires that one accepts the reported exposure and disease information in the animal studies as true, even though methodological issues, including chronic respiratory infections in the animals could have biased the results. Furthermore, because of the weak and inconsistent toxicological evidence for a causal association, this account requires that one rely heavily on the truth of toxicokinetic and mechanistic hypotheses that permit a plausible biological mode of action. To accept this account as true, one must accept either of two scenarios based on the two proposed modes of action. For the first, one must accept that somehow typical environmental exposures to methanol can lead to concentrations of formaldehyde in tissues above endogenous levels, and that this increased level of formaldehyde can lead to carcinogenesis (particularly lymphomas), even though there are several lines of evidence that counter this hypothesis, including: (a) a large body of evidence suggesting the biological implausibility of increased blood and tissue levels of formaldehyde following typical human methanol exposures; (b) genotoxic and carcinogenic effects of formaldehyde are different from those observed from methanol (i.e., methanol is not genotoxic even at very high exposure concentrations, and reported carcinogenic effects from methanol are lymphomas and not leukemias that are the predominant tumor type reported in the formaldehyde studies); (c) the scientific basis for a causal association between formaldehyde exposure and leukemogenesis is weak (Cruzan, 2009; Golden et al., 2006; Pyatt et al., 2008; Rhomberg et al., 2011). For the second scenario, one must accept that the oxidative stress mode of action proposed for animal carcinogenesis is somehow relevant to humans, even though methanol is metabolized to formaldehyde via ADH1 in humans and not catalase; therefore, not producing hydrogen peroxide and oxidative stress in humans. If one is to conclude that methanol exposure is causally associated with human lymphoma, one must assert not only that these hypothesized modes of action are conceivably true but that it is indeed known that one of them is true, for otherwise an essential and necessary element of the causal conclusion is missing. For this account (exposure to methanol is causally associated with lymphoma in humans), there is a very large degree of ad hoc argument. That is, the elements of this account are chosen so as to fit the hypothesis already put forth, not based purely on an evaluation of the weight of the evidence as a whole and how it may (or may not) support the proposed hypothesis. Consequently, alternative accounts need to be considered.

An alternative, and contrasting, account is that it is not possible for methanol to increase levels of formaldehyde in the blood and tissues from typical human methanol exposures, or for the metabolism of methanol in humans to lead to a condition of oxidative stress, and therefore there is no biologically plausible mechanism for methanol carcinogenesis in humans. This is supported by a large body of toxicokinetic and mechanistic data in animals and in vitro, and by the largely negative toxicological evidence, which are considered under this account to be the true results, while those relatively isolated positive results are considered false positives attributable to potential flaws in experimental design and confounding by chronic respiratory infection. If this account is true, an association between methanol exposure and carcinogenesis (lymphoma) would be understood as not plausible for humans, and the few positive associations that have been observed in animals would be attributed to alternative explanations. That is, the apparent positive responses among the animal studies may actually have been some combination of false positives attributable to possible methodological problems and chronic respiratory infections in the study animals. Or, given the evidence suggesting that the formaldehyde mode of action pathway is biologically not plausible, and the tumor types reported in the methanol-exposed animals were not the same as those reported in the formaldehyde studies, the results in the animal studies (if found to be truly related to methanol exposure) could only be due to a mechanism not relevant in humans (oxidative stress from catalase metabolism of methanol).

In comparing these two accounts, neither is proven or disproven, but when assessing the weight of the available evidence in support of either account, it is clear that the first account requires far more ad hoc assumptions and post hoc explanations. In the first account, the inferences regarding potential human risk do not come from the data themselves, but from assumptions invoked after the fact to fit the hypotheses put forth and without the evidence that would tie the weak toxicological and mode of action data causally to methanol exposure. Therefore, the weight of evidence for this account (i.e., exposure to methanol is causally associated with lymphoma in humans) is weak in comparison to the more substantial weight of evidence supporting the lack of a causal association.

Conflict of interest statement

The authors' affiliation is as shown on the first page. This paper was prepared with financial support to Gradient, a private environmental consulting firm, from the Methanol Institute, a trade association representing producers and distributors of methanol products. The work reported in the paper was conducted during the normal course of employment by Gradient. The authors have the sole responsibility for the writing and contents of this paper. The views and opinions expressed are not necessarily those of the Methanol Institute.

References

Belpoggi, F., Soffritti, M., Maltoni, C., 1995. Methyl-tertiary-butyl ether (MTBE) – a gasoline additive – causes testicular and lympho-haematopoietic cancers in rats. Toxicol. Ind. Health 11, 119–149.

Belpoggi, F., Soffritti, M., Maltoni, C., 1999. Immunoblastic lymphomas in Sprague-Dawley rats following exposure to the gasoline oxygenated additives methyl-tertiary-butyl ether (MTBE) and ethyl-tertiary-butyl ether (ETBE): early observations on their natural history. Eur. J. Oncol. 4, 563–572.

Bouchard, M., Brunet, R.C., Droz, P.O., Carrier, G., 2001. A biologically based dynamic model for predicting the disposition of methanol and its metabolites in animals and humans. Toxicol. Sci. 64, 169–184.

Caldwell, J.C., Jinot, J., DeVoney, D., Gift, J.S., 2008. Evaluation of evidence for infection as a mode of action for induction of rat lymphoma. Environ. Mol. Mutagen. 49, 155–164.

Chang, L.W., McMillan, L., Wynne, B.R., Pereira, M.A., Colley, R.A., Ward, J.B., Legator, M.S., 1983. The evaluation of six different monitors for the exposure to formaldehyde in laboratory rats. Mol. Carcinog. 5, 381 (in: Abstracts of the Fourteenth Annual Meeting of the Environmental Mutagen Society, March 3–6, 1983, San Antonio, Texas, abstract Bb-13.).

Cruzan, G., 2009. Assessment of the cancer potential of methanol. Crit. Rev. Toxicol. 39, 347–363.

Cruzan, G., 2011. Evaluation of the Ramazzini Foundation Study of methanol in rats: a comparison of diagnoses by the RF study pathologist and a recent NTP review team. Submitted to Methanol Institute. Available from <http://www.epa.gov/quality/informationguidelines/documents/10005-rel2.pdf>.

EFSA (European Food Safety Authority), 2006. Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food (AFC) on a request from the Commission related to a new long-term carcinogenicity study on aspartame: question number EFSA-Q-2005-122. EFSA J. 356, 1–44.

ERG (Eastern Research Group, Inc.), 2009. External letter peer review of reports documenting methanol studies in monkeys, rats and mice performed by the New Energy Development Organization (NEDO): peer reviewer comments. Submitted to US EPA, National Center for Environmental Assessment. Available from <http://www.epa.gov/ncea/pdfs/iris/meth/TO64_Methanol_Review_Deliverable_NEDO_LetterPeerReview.pdf>.

Golden, R., Pyatt, D., Shields, P.G., 2006. Formaldehyde as a potential human leukemogen: an assessment of biological plausibility. Crit. Rev. Toxicol. 36, 135–153.

Gul, H., Lu, K., Upton, P., Swenberg, J.A., 2011. Formaldehyde-induced hydroxymethyl DNA adducts in rats exposed to isotope labeled methanol. Toxicologist 120, 364 (abstract 1691).

Hayes, A.W., 2001. Principles and Methods of Toxicology, fourth ed. Taylor & Francis, Philadelphia, PA.

Hill, A.B., 1965. The environment and disease: association or causation? Proc. R. Soc. Med. 58, 295–300.

Horton, V.L., Higuchi, M.A., Rickert, D.E., 1992. Physiologically based pharmacokinetic model for methanol in rats, monkeys, and humans. Toxicol. Appl. Pharmacol. 117, 26–36.

Kavet, R., Nauss, K.M., 1990. The toxicity of inhaled methanol vapors. Crit. Rev. Toxicol. 21, 21–50.

Lu, K., Collins, L.B., Ru, H., Bermudez, E., Swenberg, J.A., 2010. Distribution of DNA adducts caused by inhaled formaldehyde is consistent with induction of nasal carcinoma but not leukemia. Toxicol. Sci. 116, 441–451.

Lu, K., Moeller, B., Doyle-Eisele, M., McDonald, J., Swenberg, J.A., 2011. Molecular dosimetry of *N*(2)-hydroxymethyl-dG DNA adducts in rats exposed to formaldehyde. Chem. Res. Toxicol. 24, 159–161.

Malarkey, D., Herbert, R., 2010. Unpublished internal memo to J. Bucher re: report on visit (4/25/2010–4/30/2010) and assessment of the pathology procedures performed at the Ramazzini Institute (RI), Bentivoglio, Italy. National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC.

McCallum, G.P., Sui, M., Ondovcik, S.L., Sweeting, J.N., Wells, P.G., 2011a. Methanol exposure does not lead to accumulation of oxidative DNA damage in bone marrow and spleen of mice, rabbits or primates. Mol. Carcinog. 50, 163–172.

McCallum, G.P., Siu, M., Sweeting, J.N., Wells, P.G., 2011b. Methanol exposure does not produce oxidatively damaged DNA in lung, liver or kidney of adult mice, rabbits or primates. Toxicol. Appl. Pharmacol. 250, 147–153.

MI (Methanol Institute), 2010. Comments on the U.S. EPA Draft Toxicological Review of Methanol (IRIS). Submitted to US EPA Docket, EPA-HQ-ORD-2009-0398-0019.1.

Moeller, B.C., Lu, K., Doyle-Eisele, M., McDonald, J., Gigliotti, A., Swenberg, J.A., 2011. Determination of *N*(2)-hydroxymethyl-dG adducts in the nasal epithelium and bone marrow of nonhuman primates following (13)CD(2)-formaldehyde inhalation exposure. Chem. Res. Toxicol. 24, 162–164.

NIOSH (National Institute for Occupational Safety and Health), 2007. NIOSH pocket guide to chemical hazards. DHHS (NIOSH) publication no. 2005-149. Available from <http://www.cdc.gov/niosh/npg/>.

NRC (National Research Council), 2011. Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde. National Academies Press, Washington, DC. Available from <http://www.nap.edu/catalog/13142.html>.

NTP (National Toxicology Program), 2010. 2-Year study protocol. Available from <http://ntp.niehs.nih.gov/?objectid=36305D16-F1F6-975E-79776DAD38EC101E>.

OECD (Organisation for Economic Co-operation and Development), 2009. OECD test guideline 453: combined chronic toxicity/carcinogenicity studies. OECD 453. Available from <http://www.oecdilibrary.org/oecd/content/book/9789264071223-en>.

Parthasarathy, N.J., Kumar, R.S., Manikandan, S., Devi, R.S., 2006. Methanol-induced oxidative stress in rat lymphoid organs. J. Occup. Health 48, 20–27.

Pereira, M.A., Chang, L.W., McMillan, L., Ward, J.B., Legator, M.S., 1982. Battery of short-term tests in laboratory animals to corroborate the detection of human population exposures to genotoxic chemicals. Mol. Carcinog. 4, 317 (in: Abstracts of the Thirteenth Annual Meeting of the Environmental Mutagen Society, February 26–March 1, 1982, Boston, Massachusetts).

Prueitt, R.L., Goodman, J.E., Bailey, L.A., Rhomberg, L.R., 2011. Hypothesis-Based Weight-of-Evidence evaluation of the neurodevelopmental effects of chlorpyrifos. Crit. Rev. Toxicol. 41, 822–897.

Pyatt, D., Natelson, E., Golden, R., 2008. Is inhalation exposure to formaldehyde a biologically plausible cause of lymphohematopoietic malignancies? Regul. Toxicol. Pharmacol. 51, 119–133.

Rhomberg, L.R., Bailey, L.A., Goodman, J.E., 2010. Hypothesis-Based Weight-of-Evidence: a tool for evaluating and communicating uncertainties and inconsistencies in the large body of evidence in proposing a carcinogenic mode of action – Naphthalene as an example. Crit. Rev. Toxicol. 40, 671–696.

Rhomberg, L.R., Bailey, L.A., Goodman, J.E., Hamade, A., Mayfield, D., 2011. Is exposure to formaldehyde in air causally associated with leukemia? – A Hypothesis-Based Weight-of-Evidence analysis. Crit. Rev. Toxicol. 41, 555–621.

Schoeb, T.R., McConnell, E.E., 2011a. Commentary: further comments on *Mycoplasma pulmonis* and lymphoma in bioassays of rats. Vet. Pathol. 48, 420–426.

Schoeb, T.R., McConnell, E.E., 2011b. *Mycoplasma pulmonis* and lymphoma in a methanol bioassay. Vet. Pathol. 48, 903–905.

Schoeb, T.R., McConnell, E.E., Juliana, M.M., Davis, J.K., Davidson, M.K., Lindsey, J.R., 2009a. *Mycoplasma pulmonis* and lymphoma in bioassays in rats. Vet. Pathol. 46, 952–959.

Schoeb, T.R., McConnell, E.E., Juliana, M.M., Davis, J.K., Davidson, M.K., Lindsey, J.R., 2009b. *Mycoplasma pulmonis* and lymphomas (letter). Environ. Mol. Mutagen. 50, 1–3.

Soffritti, M., Belpoggi, F., Cevolani, D., Guarino, M., Padovani, M., Maltoni, C., 2002. Results of long-term experimental studies on the carcinogenicity of methyl alcohol and ethyl alcohol in rats. Ann. N. Y. Acad. Sci. 982, 46–69.

TERA (Toxicology Excellence for Risk Assessment), 2010. Report on the peer review of the methanol bioassays: Soffritti et al. (2002) and Apaja (1980). Report to Methanol Institute. Available from <http://www.tera.org/peer/MeetingReports/MethanolFinalReportRevised.pdf>.

US EPA, 2009. Toxicological review of methanol in support of summary information on the Integrated Risk Information System (IRIS) (external peer review draft). EPA/635/R-09/013. Available from<http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=56521>.

US EPA, 2010. Toxicological review of formaldehyde - inhalation assessment (CAS No. 50-00-0) in support of summary information on the Integrated Risk Information System (IRIS), vol. I–IV (draft). EPA/635/R-10/002A.

Wang, M., Cheng, G., Villalta, P., Kassie, F., Hecht, S., 2008. Analysis of formaldehyde-DNA adducts in leukocyte DNA by liquid chromatography-electrospray ionization-tandem mass spectrometry: application to rats treated with aspartame [abstract]. Abstract No. 2250, in 99th AACR Annual Meeting, April 12-16, 2008, San Diego, CA. Available from <http://www.aacrmeetingabstracts.org/cgi/content/meeting_abstract/2008/1_Annual_Meeting/2250?maxtoshow=&hits=10&RESULTFORMAT=&author1=wang&fulltext=formaldehyde&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT>.